Neutron-diffraction studies of the structure of biological macromolecules

Yu. M. Ostanevich and I. N. Serdyuk

Joint Institute for Nuclear Research, Dubna (Moscow District) and Protein Institute of the Academy of Sciences of the USSR, Pushchino (Moscow District)
Usp. Fiz. Nauk **137**, 85–116 (May 1982)

Theoretical and experimental studies of the past decade on the application of thermal neutrons for studying the properties of biological macromolecules are reviewed. The physical bases are presented for the application of thermal neutrons for studying biological macromolecules. The fundamental approaches to structural problems based on a change in the isotopic composition of the solvent are discussed in detail. Results are presented of the experimental studies with solutions and monocrystals of macromolecules. The fundamental trends in further development of these studies are discussed.

PACS numbers: 87.15.By, 87.15.Da, 87.80. + s

CONTENTS

1. INTRODUCTION

The year 1982 marks the passage of fifty years since the discovery of the neutron. Among the multitude of well-known fundamental and applied consequences of this discovery, one of the first was the rise of structural neutron-diffraction analysis, one of the methods of studying the spatial structure of the condensed state of matter at the atomic level. Although this article is devoted to reviewing the novelties of the past decade in a rather narrow field—structural neutron-diffraction analysis of biological macromolecules—we deem it appropriate to present a short list of the events that have to some degree characterized the development of structural neutron-diffraction analysis as a whole.

1920: Rutherford advances the hypothesis of the existence of the neutron.

1932: J. Chadwick proves that the unusually strongly penetrating radiation discovered two years earlier by W. Bothe and H. Becker in nuclear reactions in beryllium consists of particles of mass 1 and charge 0, i.e., neutrons.

1934: E. Fermi, E. Amaldi, B. Pontecorvo, F. Rasetti, and E. Segrè discover the phenomenon of slowing of neutrons and obtain thermal neutrons.

1936: W. M. Elasasser carries out a theoretical treatment of neutron diffraction in polycrystals. H. Halban and P. Preiswerk obtain an experimental proof of neutron diffraction in polycrystalline iron, and D. P. Mitchell and P. N. Powers in monocrystals of MgO.

1947: E. Fermi and L. Marshall employ Bragg reflection of different orders and total reflection of neutrons from mirrors to determine the magnitude and sign of the coherent amplitude of scattering for 22 chemical elements.

1948: G. E. Bacon and R. D. Lowde solve the problem of taking account of secondary extinction in large monocrystals.

1951: C. G. Shull and E. O. Wollan publish a table of the coherent scattering amplitudes for 60 nuclei.

1952: S. W. Peterson and H. A. Levy carry out a complete neutron-diffraction study of the structure of a monocrystal of KHF_2.

 0038-5670/82/050323-17$01.80 © 1982 American Institute of Physics

1964: J. M. Brown and H. A. Levy carry out a complete neutron-diffraction study of the structure of sucrose $C_{12}H_{22}O_{11}$,

1967: D. C. Hodgkin, F. M. Moore, and B. T. M. Willis perform a neutron-diffraction study of a monocrystal of vitamin B_{12} ($C_{63}H_{87}O_{15}N_{13}PCo\cdot 16H_2O$). The coordinates of 205 atoms, including 98 hydrogen atoms, are established.

1969: J. Schelten, A. Mayer, W. Schmatz, and F. Hossfeld study the small-angle scattering of neutrons in solutions of biological macromolecules. B. P. Schoenborn undertakes a neutron-diffraction study of a single crystal of myoglobin.

1972: The table of coherent scattering amplitudes includes values for 157 nuclei.

1981: The number of studies macromolecular systems in solutions is of the order of 100, and studies are being performed on at least five single crystals of biological macromolecules. The quaternary structure of ribosomes is studied.

It is evident from this certainly incomplete list that the advent of the "biological period" is a fully regular stage in the evolution of structural neutron-diffraction investigations.

The neutron-diffraction methods of studying the properties of matter rest on the properties of the neutron as an elementary particle, of which the most important are the lack of electric charge, the existence of a rest mass, a rather strong interaction with atomic nuclei and a weak interaction with diamagnetic electrons. The aggregate of these properties makes neutrons a strongly penetrating form of radiation whose wave properties can be elicited by adequate dimensions of the objects of study (usually in the range 1–10 Å), while the kinematic properties (velocity, kinetic energy) are convenient for studying processes of relative motion of atoms and molecules in matter. There is a deep analogy between neutron optics and the optics of light and x-rays,[1] which is sharply manifested in the diffraction methods of studying the structure of matter.

At the same time, neutron scattering has its own specifics, which in a great number of cases renders the application of neutrons the more favored or the only suitable method of study. Precisely this specifics has given rise to the considerable advances of structural neutron-diffraction study for studying biological objects.

2. BRIEF INFORMATION ON THE INTERACTION OF THERMAL NEUTRONS WITH NUCLEI

Like other elementary particles, neutrons possess wave properties. For structural studies of biological objects at the molecular level, the most suitable are the so-called thermal neutrons, i.e., neutrons having a kinetic energy comparable with that of a gas at room temperature. As the de Broglie equation implies, the wavelength of such neutrons is about 1.5 Å. The maximal flux of neutrons in reactors utilizing light or heavy water at room temperature as the moderator is found precisely in this range of wavelengths. This range of wavelengths is optimal for determining the coordinates of atoms, since the distance between atoms is of the same order of magnitude. A cold moderator (liquid deuterium at a temperature about 20 K) shifts the maximum of the neutron flux toward longer wavelengths (4–6 Å) These wavelengths are optimal for most small-angle neutron spectrometers designed to study biological objects in solution by diffuse scattering.

Neutrons are scattered by the nuclei of atoms. (A special literature is devoted to describing the details of this process,[2] and hence we shall restrict the treatment to the minimum of information.) The dimensions of nuclei are of the order of 10^{-12} cm, which is four orders of magnitude smaller than the wavelength of thermal neutrons. Consequently the amplitude of the scattered neutron wave is isotropic, real, and practically independent of the energy of the neutron for all the major atoms contained in biopolymers. However, the sign and magnitude of the scattering amplitude depend in irregular fashion on the charge and mass number of the nucleus, as well as on the mutual orientation of the spins of the neutron and the nucleus. For convenience of description of the properties of simple ensembles of atoms (e.g., pure elements or their isotopes), one introduces two characteristics for each of them: the coherent and incoherent scattering amplitudes, which are defined by the following relationships:

$$b_{\text{coh}} = \langle b_i \rangle, \tag{2.1}$$

$$b_{\text{inc}}^2 = \langle b_i^2 \rangle - \langle b_i \rangle^2. \tag{2.2}$$

Here $\langle \ldots_i \rangle$ denotes averaging over all the values of the "pure" scattering amplitudes with weights that reflect the isotopic composition and the statistical weight of the spin states. The meaning of this separation of the averaged quantities becomes clear when we examine the expressions for the intensity of coherent and incoherent scattering of a unit flux of neutrons by some sufficiently small ($V \ll \lambda^3$) ensemble of atoms:

$$I_{\text{coh}} = 4\pi N^2 b_{\text{coh}}^2, \tag{2.3}$$

$$I_{\text{inc}} = 4\pi N b_{\text{inc}}^2. \tag{2.4}$$

The first of these is proportional to the square of the number of particles in the ensemble, which indicates the interference of the waves scattered by all the particles. The coherent component of the intensity is structure-sensitive. The second expression, which is proportional to N, does not possess this property. In structural experiments it reflects the level of the isotropic background involving the non-identity of the scattering centers. For x-rays there is no need to introduce separately the coherent and incoherent components. Hence, for description one usually employs the so-called atomic form factor f, which formally coincides with b_{coh} for neutrons.

Table I presents the fundamental parameters of neutron and x-ray scattering for the atoms of greatest interest. The data for the "heavy" atom uranium are presented here also. Henceforth we shall use b instead of b_{coh}.

Several important conclusions stem from the data of Table I.

TABLE I. Fundamental scattering characteristics for neutrons and x-rays of certain atoms.

Atom	Neutrons				X-rays
	b_{coh}	σ_{coh}	σ_s	σ_a	$f(\theta=0)$
^{1}H	−0.374	1.8	81.6	0.19	0.28
^{2}H (D)	+0.667	5.6	7.6	0.0005	0.28
^{12}C	+0.665	5.55	5.6	0.003	1.69
^{14}N	+0.940	11.1	11.1	1.1	1.97
^{16}O	+0.580	4.2	4.2	0.0001	2.25
^{31}P	+0.51	3.3	3.6	0.09	4.23
^{32}S	+0.285	1.0	1.2	0.28	4.50
U	+0.85	9.1	9.2	7.68	25.90

Note. The values of b_{coh} and f are given in units of 10^{-4} m. The coherent scattering cross-section (σ_{coh}), the total scattering cross-section (coherent + incoherent) (σ_s) and the absorption cross-section (σ_a) are given in units of 10^{-28} m^2 (b). The values of b_{coh} are taken from Ref. 3.

The scattering amplitude for neutrons depends irregularly on the atomic number. The scattering amplitudes for the atoms C, N, and O are of the same order of magnitude as for the heavy metals. Thus neutron scattering essentially differs from x-ray scattering, in which the scattering amplitude is directly proportional to the atomic number. For example, the x-ray scattering amplitude for uranium is 15 times larger than for carbon, and 12 times larger than for nitrogen. We note that the x-ray scattering amplitude depends strongly on the scattering angle, since the dimensions of the electron cloud are close to the wavelength of the x-ray radiation (1.5 Å) In Table I the x-ray scattering amplitudes f are given for zero scattering angle.

The neutron scattering amplitudes for the atoms of hydrogen and deuterium differ in magnitude and in sign. Thus the hydrogen atom will have a negative density in neutron scattering-amplitude density maps, whereas all the other atoms from Table I will have a positive density. The isotope substitution H–D enables one to "control" the scattering amplitude.

The incoherent scattering cross-section is small in comparison with the coherent cross-section for all atoms except the hydrogen atom. Hence deuteration of hydrogen-containing materials will substantially diminish the background scattering.

3. ORDER AND STRUCTURAL STUDIES

In the cases of interest to us, the amplitude of the elastically and coherently scattered wave can be written as a Fourier transform[4]:

$$A(\varkappa) = \int_V \rho(\mathbf{r}) \exp(i\varkappa\cdot\mathbf{r})\, dV. \tag{3.1}$$

Here $\rho(\mathbf{r})$ is the spatial distribution of the coherent scattering-amplitude density. The integration is performed over the whole volume V of the studied specimen, while the Fourier transform creates an "image" of the studied system in $\varkappa$-space, where

$$\varkappa = \mathbf{k} - \mathbf{k}_0 \text{ is the scattering vector,} \tag{3.2}$$

$$\varkappa = 4\pi \sin(\theta/2)/\lambda. \tag{3.3}$$

Here λ is the de Broglie wavelength and θ is the scattering angle. The experimentally observable quantity I—the neutron-scattering intensity—is proportional to

$$I(\varkappa) = |A(\varkappa)|^2 = \left| \int_V \rho(\mathbf{r})\, e^{i\varkappa\mathbf{r}}\, dV \right|^2. \tag{3.4}$$

The task of structural studies is to reconstruct from the measured $I(\varkappa)$ relationship the spatial density distribution $\rho(\mathbf{r})$ in the object of study. The degree of solvability of this problem depends on a large number of factors, among which the periodicity of the studied structure plays the leading role. One can attain the most complete solution if the studied object is a monocrystal with a high degree of perfection, consisting of identical macromolecules. Then it proves possible to reconstruct $\rho(\mathbf{r})$ with a spatial resolution of the order of the dimensions of the individual atoms. The attainable spatial resolution sharply deteriorates as the studied systems become disordered. For dilute solutions of macromolecules, it proves to be of the order of the dimensions of the molecules themselves, and one can decide on their structures only in special cases. Both extreme cases have been studied in sufficient detail methodologically and are widely applied in structural studies. Of course, one must remember in evaluating structural studies that their final aim is to solve biological problems on the level of structural organization that appears adequate to the problems that have been posed. The large number of structural levels in molecular biology creates wide possibilities for both the high- and low-resolution methods.

4. SMALL-ANGLE DIFFUSE SCATTERING OF NEUTRONS

Small-angle diffuse scattering is an example of the methods of structural analysis at low spatial resolution.[5] The integration in the formulas (3.1) and (3.4) is restricted to the volume of one macromolecule, which implies infinitely dilute systems. We shall briefly take up the fundamental approaches currently being applied in small-angle neutron scattering, since there are as yet no reviews or monographs on this topic in the literature of our country.

We can derive the expression (3.4) for a single macromolecule in a solution (differential scattering cross-section) by averaging (3.4) over the orientation. For isotropic solutions the averaging yields the result:

$$\left\langle \frac{d\sigma}{d\Omega} \right\rangle = \int_V \int_{V'} \rho(\mathbf{r})\, \rho(\mathbf{r}') \frac{\sin \varkappa |\mathbf{r}-\mathbf{r}'|}{\varkappa |\mathbf{r}-\mathbf{r}'|}\, dV\, dV'. \tag{4.1}$$

This is called the Debye formula.

The intensity of the scattered neutrons is an even function of the length $\varkappa$ of the scattering vector, as is natural for an isotropic scatterer. If the macromolecules have a characteristic dimension L, the expression (4.1) directly yields some fundamental properties of the scattering law. When $\varkappa L \ll 1$, the factor $\sin x/x$ becomes unity, the scattering intensity is $\sim \rho^2 V^2$, and the latter depends neither on the internal structure nor the shape of the macromolecule. When $\varkappa L \approx 1$, the intensity fundamentally depends on the dimensions and shape of the macromolecule. Then $\varkappa L \gg 1$, in principle the intensity becomes sensitive to the details of internal structure of the macromolecule, but the oscillating character of the integrand leads to a rapid decay of the

intensity with increasing $\varkappa$ [in the general case ~$1/(\varkappa L)$].[4] The major part of the scattered intensity lies in the region $\varkappa L \lesssim 1$. For typical dimensions of the macromolecule of 50 Å and with a neutron wavelength of 5 Å, this corresponds to $\sin(\theta/2) \leqslant \lambda/4\pi L = 8 \times 10^{-3}$ or to angles $\leqslant 1°$. The latter situation explains the source of the term "small-angle scattering".

The discussion above has ignored the properties of the solvent. If we treat the solvent as a homogeneous, isotropic medium, we can characterize it with the constant scattering density

$$\rho_s = \frac{1}{V}\sum_{i\in V} b_i. \tag{4.2}$$

Here V is some sufficiently large volume (e.g., 1 cm^3), while the sum of the coherent amplitudes is taken over all the atoms contained in this volume. All the previous arguments remain in force if we use the difference of scattering densities $\rho(\mathbf{r}) - \rho_s$ instead of $\rho(\mathbf{r})$.

a) Homogeneous particles

An important approximation for many problems in the approximation of the homogeneous particle: $\rho(\mathbf{r}) = \text{const} = \bar{\rho}$. Then the expression for the differential scattering cross-section for a single particle [Eq. (4.1)] is transformed into the form

$$\left\langle \frac{d\sigma}{d\Omega} \right\rangle = (\bar{\rho} - \rho_s)^2 \left\langle \left| \int_V e^{i\varkappa \mathbf{r}}\, dV \right|^2 \right\rangle = (\bar{\rho} - \rho_s)^2 V^2 F^2(\varkappa). \tag{4.3}$$

The function $F^2(\varkappa)$ depends only on the shape and dimensions of the particle being studied; in some cases it can be calculated analytically.[5] For example, for a sphere of radius R we have

$$F^2(\varkappa) = \left(\frac{3\sin \varkappa R - \varkappa R \cos \varkappa R}{\varkappa^3 R^3} \right)^2. \tag{4.4}$$

For an ellipsoid of revolution with semiaxes a, a, and Pa we have

$$F^2(\varkappa) = \int_0^{\pi/2} \Phi^2\left(\varkappa a \sqrt{\cos^2\theta + P^2 \sin^2\theta}\right) \cos\theta\, d\theta. \tag{4.5}$$

Here Φ is the error integral. For an infinitely thin disk of radius R we have

$$F^2(\varkappa) = \frac{2}{\varkappa^2 R^2}\left[1 - \frac{1}{\varkappa R} J_1(2\varkappa R)\right]. \tag{4.6}$$

For an infintely thin rod of length $2H$ we have

$$F^2(\varkappa) = \frac{\mathrm{Si}(2\varkappa H)}{\varkappa H} - \frac{\sin^2(\varkappa H)}{\varkappa^2 H^2}. \tag{4.7}$$

For a flexible chain molecule with a mean-square distance between the ends of $\bar{h}^2$, we have[6]

$$F^2(\varkappa) = 2\left(e^{-x} + x - 1\right) x^{-2}. \tag{4.8}$$

Here we have $x = \varkappa^2 \bar{h}^2/6$.

We see from Eqs. (4.2)–(4.8) that the argument of the functions is always the dimensionless quantity $\varkappa L$, where L is some dimension of the particle. This implies that, when one changes this dimension by a factor of α (scale transformation), the scattering law can be obtained by a simple change of the scale of $\varkappa$. In other words, similar objects have similar scattering indicatrices, or $F^2(\varkappa, \alpha L) = F^2(\alpha\varkappa, L)$. This circumstance makes it rather attractive to calculate numerically $F^2(\varkappa)$ for objects of defined shapes, represented as functions of the dimensionless argument. If one can attain agreement of the experimental with the calculated scattering curve by a simple transformation of the argument, it becomes desirable to draw conclusions on both the dimensions and the shape of the macromolecule. This approach, which is called the method of standard curves, has been developed by Kratky and his associates.[7] However, one must apply this method with great circumspection, since this approach is based on the hypothesis of homogeneity of the studied object. Also, most unpleasantly, the scattering curves can coincide fortuitously for objects of completely different structures. Thus, for example, the scattering curve for a Gaussian coil is practically indistinguishable over a region of decline by 1.5 orders of magnitude from that for a homogeneous prolate ellipsoid of revolution with an axial ratio of 5:1. Practically useful methods of calculating $F^2(\varkappa)$ for objects of complex shape have been developed by Feĭgin and his associates.[8]

Returning to Eq. (4.2), we can easily see that the scattering intensity for a homogeneous particle vanishes for all $\varkappa$ if $\bar{\rho} = \rho_s$. For neutrons this situation is easily realizable by choosing a suitable isotopic composition of the solvent (water). For light and x-rays the condition $\bar{\rho} = \rho_s$ can be fulfilled in some cases by choosing the chemical composition of the solvent. This involves a substantially large risk of altering the structure of the macromolecule being studied. Variation of the scattering indicatrix by varying ρ_s has been called the method of contrast variation. The structural aspects that can be treated in the following sections.

b) Inhomogeneous particles. General approach

A number of problems important in molecular biology involve the study of *inhomogeneous particles*. The "inhomogeneities" must be coarse enough—of superatomic dimensions, while the physical nature of the inhomogeneities can be anything whatever (chemical composition, isotopic composition, etc.), as long as the neutron-optic characteristics vary The most general approach to this problem is associated with the studies of Stuhrmann and Kirste,[9] who proposed to represent $\rho(\mathbf{r})$ in the form of two components

$$\rho(\mathbf{r}) = \bar{\rho} + \rho_F(\mathbf{r}). \tag{4.9}$$

Here $\rho_F(\mathbf{r})$ is the fluctuating component of the density, which is defined by the condition

$$\int_V \rho_F(\mathbf{r})\, dV = 0. \tag{4.10}$$

Further, the mean density is

$$\bar{\rho} = \frac{1}{V}\int_V \rho(\mathbf{r})\, dV. \tag{4.11}$$

When this representation is taken into account, the scattering cross-section acquires the form of the sum of three terms:

$$\left\langle \frac{d\sigma}{d\Omega} \right\rangle = \left\langle \left| (\bar{\rho} - \rho_s) \int_V e^{i\varkappa r} dV + \int_V \rho_F(\mathbf{r})\, e^{i\varkappa r} dV \right|^2 \right\rangle$$
$$= (\bar{\rho} - \rho_s)^2 F_V^2(\varkappa) + (\bar{\rho} - \rho_s) F_{VF}(\varkappa) + F_F^2(\varkappa); \quad (4.12)$$

Here F_V^2 depends only on the shape and volume of the particle under study ("homogeneous" particle), and F_F^2 depends only on the fluctuations of the scattering density about the mean $\bar{\rho}$, while the middle term F_{FV} is of mixed nature. The possibility of separating the components F_V^2 and F_F^2 is important. Evidently one can measure $F_F^2(\varkappa)$ at the compensation point $(\bar{\rho} - \rho_s = 0)$. To measure $F_V^2(\varkappa)$, one extrapolates to "infinite contrast" $[1/(\bar{\rho} - \rho_s) = 0]$. Finally, one can analyze the quadratic form (4.12) as a function of $\bar{\rho} - \rho_s$, while seeking all three unknown F's for each $\varkappa$.

The approach presented here looks very attractive, but it has a yet not found widespread application in its full scope. In performance it faces two serious difficulties. The first of them involves the fact that deuteration of the solvent (water) causes a fraction of the protons in the studied particle (macromolecule) also to be replaced by deuterons. Consequently $\rho(\mathbf{r})$ becomes a function of a large number of new variables that describe the capability for interchange and the spatial distribution of the exchangeable protons. In particular, one can study the pattern of density fluctuations only at the compensation point, though here we shall see fluctuations somewhat different from those far from this point. In this regard F_V^2 is less sensitive to H–D exchange, owing to the fact that the extrapolation to infinite contrast is performed from points of sufficiently high contrast [the first term in (4.12) predominates]. The second difficulty involves the fact that it is difficult to draw any at all detailed conclusion concerning $\rho(\mathbf{r})$ from the measured F_V^2 and F_F^2 (see Sec. 4g).

c) Scattering in the very-small-angle region, radius of gyration, Guinier approximation

In 1939 Guinier took an important step on the path of converting Eq. (4.1) into a working instrument.[5] Upon employing the expansion $\sin x/x = 1 - x^2/3! + x^4/5! - \ldots$ and retaining the first two terms, one can simplify Eq. (4.1) further:

$$\left\langle \frac{d\sigma}{d\Omega} \right\rangle = \left(\int_V \rho(\mathbf{r})\, dV \right)^2 \left(1 - \frac{1}{3} \varkappa^2 \int_V \rho(\mathbf{r})\, r^2\, dV \Big/ \int_V \rho(\mathbf{r})\, dV + \ldots \right)$$
$$= \left(\int_V \rho(\mathbf{r})\, dV \right)^2 (1 - \varkappa^2 R_{g_0}^2/3 + \ldots). \quad (4.13)$$

Here $R_{g_0}^2$ coincides with the definition of the square of the radius of gyration of an object having a density equal equal to the scattering density. Guinier proposed that a good approximation of the expression in the last parentheses of (4.13) is the function

$$F^2(\varkappa) = \exp(-\varkappa^2 R_g^2/3). \quad (4.14)$$

As before, the parameter R_g here is the radius of gyration (Guinier approximation). The latter expression is fundamental for determining the "observed" radius of gyration from the experimental data by drawing the so-called Guinier graph of $\ln I(\varkappa^2)$ and determining the slope of the rectilinear region:

$$R_g^2 = -\frac{1}{3} d \ln I(\varkappa^2)/d\varkappa^2. \quad (4.15)$$

For homogeneous objects of simple shape one can easily calculate $R_{g_0}^2$:

$R_{g_0}^2 = 3R_0^2/5$ — sphere of radius R_0 , (4.16)

$R_{g_0}^2 = (a^2 + b^2 + c^2)/5$ — triaxial ellipsoid with semiaxes a, b, c, (4.17)

$R_{g_0}^2 = L^2/12$ — infinitely thin rod of length L, (4.18)

$R_{g_0}^2 = \bar{h}^2/6$ — Gaussian coil with the mean-square distance $\bar{h}^2$ between the ends. (4.19)

The Guinier approximation becomes exact as $\varkappa^2 \to 0$, and can usually be applied in practice in the region $\varkappa^2 R_g^2 < 1$. The deviations of $I(\varkappa)$ from the Guinier region serve to indicate the character of the anisometry (a measure of the nonsphericity) of the particle.

For a biaxial ellipsoid with a ratio of the axes $c/a = 1.86$, these deviations are insignificant up to $\varkappa^2 R^2 = 5$; for ellipsoids of smaller anisometry and spheres we have $I(\varkappa) \leq I(0) \exp(-\varkappa^2 R_g^2/3)$; the converse is true for $c/a > 1.86$. Other approximations exist on a par with (4.14). For example, in studying biological macromolecules in the unfolded state, for which the Gaussian coil often serves as a model, one usually employs the Debye approximation.

$$I^{-1}(\varkappa) = I^{-1}(0)(1 + \varkappa^2 R_g^2/3) \quad (4.20)$$

(see, e.g., Ref. 6), which has the same limiting meaning.

In addition to the approximations that have a rigorous limiting meaning, in practice constructions are employed that lead to the so-called radii of gyration of transverse cross-sections. The corresponding approximations

$$\varkappa I(\varkappa) = \text{const} \cdot \exp(-\varkappa^2 R_{q_1}^2/2) \quad (4.21)$$

and

$$\varkappa^2 I(\varkappa) = \text{const} \cdot \exp(-\varkappa^2 R_{q_2}^2) \quad (4.22)$$

are used for highly extended and highly flattened particles, respectively.[7] The parameters $R_{q_1}^2$ and $R_{q_2}^2$ denote the square of the radius of gyration of the density distribution in a two-dimensional cross-section perpendicular to the long axis of the elongated particle and the second moment of the density distribution in thickness for the flattened particle. These parameters become asymptotically exact for infinitely elongated (or flattened) particles. One usually determines the region of applicability of (4.21) and (4.22) empirically.

d) Forward scattering. Determination of the molecular weight and "dry" volume of a particle

At the vector length $\varkappa = 0$, the scattering cross-section in (4.1) for a single particle acquires an especially simple form:

$$\left\langle \frac{d\sigma}{d\Omega} \right\rangle = (\bar{\rho} - \rho_s)^2 V^2. \quad (4.23)$$

The scattering intensity from a sufficiently dilute solution containing N particles (we neglect interparticle interference) has the form

$$I(\varkappa) = I_0 N \left\langle \frac{d\sigma}{d\Omega} \right\rangle T \Delta\Omega, \quad (4.24)$$

Here I_0 is the number of neutrons incident on the solution being studied per unit time, T is a coefficient that takes into account the attenuation of the neutron beam upon passing through the solution and the efficiency of neutron detection, and $\Delta\Omega$ is the solid angle subtended by the specimen at the neutron detector.

One cannot measure the quantity $I(0)$ directly, but always obtains it experimentally by extrapolation to $\varkappa = 0$, e.g., by using the Guinier approximation. After taking the instrumental factors into one can use the following quantity to determine the molecular weight M of the particles being studied:

$$i(0) = N(\bar{\rho} - \rho_s)^2 V^2. \quad (4.25)$$

The method is based on using the independently determined partial specific volume $\bar{v}$, together with the mean scattering-amplitude density $\bar{\rho}$ in the solution. We define $\bar{v}$ by

$$\bar{v} = \frac{VN_A}{M}. \quad (4.26)$$

Here N_A is Avogadro's number. If we express N in terms of the weight concentration c of the dissolved particles (mg/ml), then we can write (4.25) in the form

$$i(0) = Mc(\bar{\rho} - \rho_s)^2 \bar{v}^2 V_s N_A^{-1}. \quad (4.27)$$

Here V_s is the volume of the solution. This expression serves for determining M.

In considering this problem for x-rays, one replaces the volume density $\bar{\rho}$ with the scattering-amplitude density per unit mass $\sum_i f_i/M = 0.15 \times 10^{-12}$ cm/dalton, which fluctuates sufficiently little already at the level of the amino-acid residues. For neutrons this representation of the density offers no special advantages, since the scattering amplitude does not depend in a regular fashion on M.

A new aspect in the neutron case is the possibility of extending this approach to multicomponent particles.[10] For such particles we have the following instead of (4.27):

$$i(0) = c\left(\sum_j (\bar{\rho}_j - \rho_s)\bar{v}_j M_j\right)^2 V_s N_A^{-1} / \sum_j M_j. \quad (4.28)$$

Here M_j, ρ_j, and $\bar{v}_j$ are the corresponding quantities for each component. Here we can determine $M = \sum_j M_j$ (the molecular weight of the particle) from measurements at one value of ρ_s if the ratios M_j/M are known independently (e.g., by chemical analysis). One can determine all the M_j's independently by measuring the dependence of $i(0)$ on ρ_s, e.g., at the compensation points of all the components, and solving the corresponding system of equations. This approach has been successfully applied to two-component systems (see Sec. 4*g*, 1). Here a certain complication is the dependence of $\bar{\rho}_j$ on ρ_s, which arises from H–D exchange. One can allow for this factor only to a certain accuracy. This will be discussed in the next section.

Another approach to (4.23) arises if we know the chemical composition of the single particle and its molecular weight. Upon determining the point ρ_{sc} at which $I(0) = 0$ from the experimental dependence of $I(0)$ on ρ_s, we have the relationship $\bar{\rho} = \rho_{sc}$, from which we can determine the so-called "dry" volume of the particle:

$$V = \sum_j h_j b_j \cdot \rho_{sc}^{-1}. \quad (4.29)$$

Here h_j is the number of atoms of the jth element in the particle, and b_j is the corresponding scattering amplitude. In calculating this summation we must also allow for H–D exchange. The dry volume of proteins determined by neutron diffraction proves to be very close to the sum of volumes of the amino-acid residues, most of which have been determined by Zamyatnin.[11] Table II contains in convenient form the information needed for calculating the summation in (4.29) from the known amino-acid or nucleotide composition.

TABLE II. Scattering amplitudes of the amino-acid residues and nucleotides calculated from their chemical composition.[74] The values in D_2O correspond to the case in which all the exchangeable hydrogen atoms in the molecule have been exchanged for deuterium with the solvent. M is the molecular weight of the residue (nucleotide), and V is its partial volume.

a) Amino-acid residues.

Amino acid	M	Chemical composition	Number of exchangeable hydrogens	$\sum h_k b_k$ (H_2O), 10^{-12} cm	$\sum h_k b_k$ (D_2O), 10^{-12} cm	V, Å^3
Glycine	57	C_2NOH_3	1	1.73	2.77	66.4
Alanine	71	C_3NOH_5	1	1.64	2.69	91.5
Valine	99	C_5NOH_9	1	1.48	2.52	141.7
Leucine	113	C_6NOH_{11}	1	1.40	2.44	167.9
Isoleucine	113	C_6NOH_{11}	1	1.40	2.44	168.8
Phenylalanine	147	C_9NOH_9	1	4.14	5.18	203.4
Tyrosine	163	$C_9NO_2H_9$	2	4.72	6.80	203.6
Tryptophan	186	$C_{11}N_2OH_{10}$	2	6.03	8.12	237.6
Aspartic acid	114	$C_4NO_3H_4$	1	3.84	4.89	113.6
Glutamic acid	128	$C_5NO_3H_6$	1	3,76	4.80	140.6
Serine	87	$C_3NO_2H_5$	2	2.22	4.31	99.1
Threonine	101	$C_4NO_2H_7$	2	2,14	4.22	122.1
Asparagine	114	$C_4N_2O_2H_6$	3	3,46	6.58	135.2
Glutamine	128	$C_5N_2O_2H_8$	3	3,73	6.50	161.1
Lysine	129	$C_6N_2OH_{13}$	4	1,59	5.75	176.2
Arginine	157	$C_6N_4OH_{13}$	6	3,47	9.71	180.8
Histidine	136,5	$C_6N_3OH_{6.5}$	1.5	4,96	6.52	167.3
Methionine	131	C_5NOSH_9	1	1,76	2.80	170.8
Cysteine	109	C_3NOSH_5	2	1,93	4.01	105.6
Proline	97	C_5NOH_7	0	2,23	2.23	129.3

b) Nucleotides.

Base		M	Chemical composition	Number of exchangeable hydrogens	$\sum h_k b_k$ (H_2O), 10^{-12} cm	$\sum h_k b_k$ (D_2O), 10^{-12} cm
Adenine	RNA	328	$PN_5C_{10}O_6H_{11}$	3	11.23	14.35
	DNA	312	$PN_5C_{10}O_6H_{11}$	2	10.65	12.73
Guanine	RNA	344	$PN_5C_{10}O_7H_{11}$	4	11.81	15.98
	DNA	328	$PN_5C_{10}O_6H_{11}$	3	11.23	14.35
Cytosine	RNA	304	$PN_3C_9O_7H_{11}$	3	9.26	12.39
	DNA	288	$PN_3C_9O_6H_{11}$	2	8.68	10.77
Uracil	RNA	305	$PN_2C_9O_8H_{10}$	2	9.28	11.36
Thymine	DNA	303	$PN_2C_{10}O_7H_{12}$	1	8.61	9.65

e) Contrast variation using H_2O-D_2O mixtures

The values of ρ for simple and composite macromolecules in H_2O as calculated by Eq. (4.11) are given in Table III. A calculation by Eq. (4.2) shows that the scattering-amplitude density of light water (-0.56×10^{10} cm^{-2}) strongly differs in sign and magnitude from that for heavy water ($+6.39 \times 10^{10}$ cm^{-2}). The scattering-amplitude density of all biological objects has an intermediate value. Hence, by putting the particle into different H_2O–D_2O mixtures one can easily attain the com-

pensation conditions, i.e., a composition of the solvent in which the scattering densities of the particle and the solvent are the same.

Deuterium exchange in the particle complicates the situation. The scattering-amplitude density ρ will depend on the fraction of D_2O in the H_2O–D_2O mixture. Generally the exact fraction of exchangeable hydrogens accessible to the solvent is unknown. If we assume that all hydrogens in NH_2–, NH–, and OH–groups are accessible to the solvent, which accounts respectively, e.g., for 23.8% and 21.4% of the total number of hydrogens in the molecules for ribonucleic acids and proteins, we can calculate the scattering-amplitude density of RNA and of a protein in D_2O. These values are 4.47×10^{-10} and 3.08×10^{-10} cm^{-2}, respectively. Upon assuming that the fraction of exchangeable hydrogens accessible to the solvent is proportional to the volume fraction (Y) of D_2O in the H_2O–D_2O mixture, we get the following relationships of $\bar{\rho}$ to Y for different biological molecules (in units of 10^{10} cm^{-2}):

$$\begin{aligned} &\text{ribosomal RNA:} && \bar{\rho}_{RNA} = +3.57 + 0.90\,Y,\\ &\text{ribosomal proteins:} && \bar{\rho}_{protein} = 1.73 + 1.35\,Y,\\ &\text{DNA:} && \bar{\rho}_{DNA} = 3.40 + 0.70\,Y,\\ &H_2O - D_2O: && \rho_s = -0.56 + 6.96\,Y. \end{aligned} \qquad (4.30)$$

Table III gives the compensation points calculated from these expressions.

A comparison of the "vacuum" densities and the densities at the compensation point with allowance for deuterium exchange shows that the latter strongly affects the scattering pattern. The assumption of complete hydrogen exchange in the polar groups gives an upper estimate of the magnitude of the exchange effects that was generally accepted until recently. However, reports have already appeared in a number of papers of a smaller degree of exchange (80% for proteins[10]). Another factor that complicates the estimate of ρ can be effects involving the presence of a layer of water of hydration and of small ions condensed on the macromolecules.[12] We must bear these considerations in mind in employing the relationships (4.30).

The radius of gyration R_g of the particle also proves to depend on the contrast if the radial scattering-density distribution is inhomogeneous or the H–D exchange proves to be inhomogeneous. For a two-component particle consisting of homogeneous components, this relationship can be written as follows[13]:

$$R_g^2 = x_1 R_1^2 + (1-x_1) R_2^2 + x_1 (1-x_1) L^2. \qquad (4.31)$$

Here R_1 and R_2 are the radii of gyration of the first and second components, L is the distance between their centers of gravity, and x_1 is the relative contribution of the first component to the total scattering amplitude:

$$x_1 = \frac{(\bar{\rho}_1 - \rho_s) V_1}{(\bar{\rho}_1 - \rho_s) V_1 + (\bar{\rho}_2 - \rho_s) V_2}; \qquad (4.32)$$

In the latter expression V_1 and V_2 are the "dry" volumes of the first and second components, which are assumed to be known independently, and the ρ_i are the mean scattering-amplitude densities of the components.

One can establish the mutual spatial arrangement of components having different scattering-amplitude densities by constructing the relationship of the square of the radius of gyration R_g^2 to the scattering fraction of one of the components. As is implied by Eq. (4.31), R_g^2 does not depend on x_1 for a homogeneous distribution of the components in a two-component particle (homogeneous particle, $R_1 = R_2, L = 0$). An inhomogeneous but symmetrical distribution of the components ($R_1 \neq R_2, L = 0$) gives rise to a linear $R_g^2(x_1)$ relationship with a nonzero slope. And finally, an inhomogeneous, asymmetric distribution of the components ($R_1 \neq R_2$ or $R_1 = R_2$ and $L \neq 0$) will be manifested in a parabolic relationship of R_g^2 to x_1.

A second method of interpreting the relation of the radius of gyration to the contrast has been proposed by Stuhrmann.[14] It uses the representation

$$R_g^2 = R_\infty^2 + \frac{\alpha}{\bar{\rho}_c - \rho_s} - \frac{\beta^2}{(\bar{\rho}_c - \rho_s)^2}. \qquad (4.33)$$

Here $\bar{\rho}_c$ is the mean scattering-amplitude density of the particle of (4.11) at the compensation point, and R_∞^2 is the radius of gyration of the particles at "infinite contrast" $[1/(\bar{\rho}_c - \rho_s) = 0]$. The constants R_∞^2, α and β have

TABLE III. Some neutron-diffraction characteristics of biological macromolecules.

Macromolecules	Experimental values: $\bar{\rho}_c$ (10^{-12}) at the compensation point	Experimental values: Compensation point, vol. % D_2O	Experimental values: Partial specific volume, cm^3/g	Calculated values: $\bar{\rho}$ (10^{-12}) in H_2O	Calculated values: $\bar{\rho}_c$ (10^{-12}) at the compensation point
a) Proteins					
1. Hemoglobin	+2.24	40^{25}	0.74_4	1.88	2.39 *)
2. Lysozyme	+2.55	45^{33}	0.70_8	1.99	2.55
3. Myoglobin	+2.24	40^{15}	0.75_1	1.87	2.32 *)
4. Fibrinogen	+2.40	43^{34}	0.72		
5. Apoferritin	+2.35	42^{61}	0.73_3	1.95	2.46
6. Ribosomal protein S4 of *E. coil*	+2.35	42^{26}	0.73_9	1.79	2.35
7. Cytochrome *c* reductase	+2.25	40^{66}	0.74		
8. EF-Tu factor of *E. coli*	+2.35	42^{70}	0.73	1.83	2.36
9. EF-Tu factor of *B. stearothermophilus*	+2.35	42^{70}	0.72_5		
10. Prothrombin	+2.35	42^{71}	0.73	1.95	2.47
b) Lipids and their complexes with proteins					
11. Phospholipid		37^{72}	0.98		−0.3
12. Low-density lipoprotein	+0.48	15^{72}	0.90		
c) RNA and its complexes with proteins					
13. Fragment of 16S ribosomal RNA	+4.32	70^{26}	0.54	3.80	4.32
14. Complex of fragment of 16S ribosomal RNA with protein S4	+3.96	65^{26}		+3.30	+3.90
15. 30S subunit of the ribosome	+3.32	56^{63}	0.60	+2.82	+3.32
16. Ditto	+3.47	58^{65}	0.60	+3.00	+3.47
17. Ditto	+3.42	57^{30}	0.60	+2.81	+3.32
18. 50S subunit of the ribosome	+3.52	5^{62}	0.59	+3.00	+3.50
19. Ditto	+3.50	58^{64}	0.59	+3.00	+3.50
20. Ditto	+3.55	59^{30}	0.59	+3.00	+3.50
21. 70S ribosome	+3.54	59^{65}	0.59	+3.00	+3.50
22. Influenza virus	+1.63	31.5^{68}	0.73		
23. Ta-t component of α-α mosaic virus	+2.64	46^{67}	0.74		
24. Type 2 adenovirus	+2.57	45^{69}	0.76		
d) DNA and its complexes with proteins					
25. DNA		$64.5^{**)\,73}$	0.51	$+3.40^{73}$	+3.94
26. Calf thymus chromatin	+2.85	49^{31}	0.61		
27. Chicken embryo erythrocyte chromatin	+2.85	49^{31}	0.62		

*Without allowing for the heme group.
**Calculated value.

no simple geometric interpretation. For a two-component particle that can be divided into two volumes V_1 and V_2 having densities $\bar{\rho}_1$ and $\bar{\rho}_2$, the constants R_∞^2, α, and β are associated with the radii of gyration of each component R_1 and R_2 and with the distance L between their centers of gravity by the following relationships:

$$R_\infty^2 = R_1^2 \frac{V_1}{V_1+V_2} + R_2^2 \frac{V_2}{V_1+V_2} + L^2 \frac{V_1 V_2}{(V_1+V_2)^2}, \quad (4.33')$$

$$\alpha = (\rho_1 - \rho_2) \frac{V_1 V_2}{(V_1+V_2)^2} \left(R_1^2 - R_2^2 + L^2 \frac{V_1 - V_2}{V_1+V_2} \right), \quad (4.34)$$

$$\beta^2 = L^2 \left[\frac{(\bar{\rho}_1 - \bar{\rho}_2) V_1 V_2}{(V_1+V_2)^2} \right]^2.$$

Special cases of the expressions (4.34) can be found in Ref. 16. An advantage of (4.33) is its applicability to inhomogeneous single-component particles. In this case one can decide from the sign of α on the character of the radial scattering-density distribution: positive values of α correspond to a particle with a "loose" core, and conversely. The parameter β differs from zero if the center of gravity of the distribution $\rho_F(\mathbf{r})$ [see (4.10)] is shifted away from the center of gravity of the particle.

The two approaches that we have discussed have been developed under the approximation of uniform deuteration, in which inhomogeneities caused by contrast variation are not considered. A mathematical apparatus has also been developed for the more complex case,[15] but has not yet found practical application.

In addition to the contrast-variation methods, which involve changing the properties of the solvent, there is another possibility of employing different types of radiations (light, x-rays, and neutrons) for studying one particular two-component particle. This possibility has been treated in Ref. 13 for synthetic block copolymers and for the fundamental two-component biological particles, and has been applied for ribosomal particles.

Among all the mentioned methods, the most widely applied is contrast variation for neutrons using H_2O–D_2O mixtures.

f) Studies of quaternary structure

In molecular biology the term "quaternary structure" denotes the mutual arrangement of the constituent parts (proteins, RNA) in composite macromolecules and their complexes. An example of this type of composite structure is the small subunit of the ribosome, which consists of one molecule of RNA and 21 different protein molecules. In this same approximation one can also treat simpler systems, e.g., enzyme-substrate complexes or individual proteins having quaternary structure, e.g., hemoglobin. In this section we shall employ the term "particle" to denote the constituent parts of systems possessing quaternary structure.

Let us consider a two-component object consisting of two identical homogeneous particles separated by the vector L. We can represent their scattering amplitude in the form

$$A = \rho \left(\int_1 e^{i\varkappa \mathbf{r}} dV + \int_2 e^{i\varkappa(\mathbf{r}+\mathbf{L})} dV \right) = \rho (1 + e^{i\varkappa \mathbf{L}}) \int e^{i\varkappa \mathbf{r}} dV.$$

The intensity is

$$I = \rho^2 \cdot 2 (1 + \cos \varkappa L) \left| \int_V e^{i\varkappa \mathbf{r}} dV \right|^2.$$

We can average over the orientations in explicit form only for spherical particles, following which the intensity contains the oscillating interference factor.

$$\langle I \rangle = 2V^2 \rho^2 F^2(\varkappa) (1 + \sin \varkappa L / \varkappa L). \quad (4.35)$$

One can determine from the observed period of the oscillations the distance between the centers of the spheres.[17]

This approach has become widely known for ribosomes, owing to the possibility of replacing individual particles, e.g., two proteins, by their copies in the fully deuterated form. In order to do this, the biological culture is grown in parallel in H_2O and D_2O media. In the D_2O medium practically all the hydrogen atoms in the ribosome are replaced by deuterium. If one possesses two subparticles (1, 2) in the fully protonated (p) and deuterated (d) forms, one can disassemble, replace the individual proteins, assemble, and obtain four types of ribosomal subparticles. The first contains both particles in the (p) form, another both particles in the (d) form, and the other two contain one deuterated particle (1 or 2) each. We can show that a combination of the four measurements

$$I_{\text{int}} = [I(1\text{p}, 2\text{p}) + I(1\text{d}, 2\text{d})] - [I(1\text{p}, 2\text{d}) + I(1\text{d}, 2\text{p})] \quad (4.36)$$

contains only an interference term depending on the distance between the centers of gravity of the two particles being studied. The attempts to use the period of the observed oscillations [see (4.35)] to determine the distance between the proteins of ribosomes have revealed a sensitivity of I_{int} to the actual shape of the particles (nonsphericity) and to their mutual orientation.[18]

A considerably more perfected approach has been developed in Ref. 19, which uses the properties of the moments of the distance distribution function.

The combination of measured $I(\varkappa)$ values that we are treating preserves only the interference term

$$I_{1,2}(\varkappa) = 2 \int_{V_1} \int_{V_2} \rho_1(\mathbf{r}_1) \rho_2(\mathbf{r}_2) \frac{\sin \varkappa |\mathbf{L} + \mathbf{r}_2 - \mathbf{r}_1|}{\varkappa |\mathbf{L} + \mathbf{r}_2 - \mathbf{r}_1|} dV_1 dV_2. \quad (4.37)$$

Here 1 and 2 refer to the particles 1 and 2, $\mathbf{r}_1$ and $\mathbf{r}_2$ are the radii from the centers of gravity of the particles, L is the vector joining their centers of gravity, and the densities ρ_1 and ρ_2 here are the differences of scattering densities in the H and D forms:

$$\rho_i(\mathbf{r}) = \rho_{i\text{H}}(\mathbf{r}) - \rho_{i\text{D}}(\mathbf{r}), \quad i = 1, 2. \quad (4.38)$$

We assume henceforth that the difference (4.38) reproduces sufficiently accurately the shape and scattering-density distribution of one particle. In particular, one can use $\rho_i(\mathbf{r})$ to calculate the radius of gyration, which has its ordinary meaning. The distribution function of the distances R between the particles 1 and 2 is related to $I(\varkappa)$ by an integral transformation[5]:

$$p(R) = R \int_0^\infty I(\varkappa) \varkappa \sin(\varkappa R) d\varkappa \quad (4.39)$$

It can be calculated if one has obtained $I(\varkappa)$ [we have $\int p(R)dR = 1$ if $\int\int \rho_1\rho_2 dV_2 dV_2 = 1$, as is assumed below]. The second moment of $p(R)$ is associated with the parameters of the particles:

$$M_2 = \int_0^\infty p(R)\, R^2\, dR = L^2 + R_{g1}^2 + R_{g2}^2. \qquad (4.40)$$

This relationship is the basis of the approach to reconstructing the quaternary structure of the small subunit of the ribosome by the method of triangulation.

The idea of the method of triangulation consists of the following. If $n \geq 3$ points are given in space, their mutual arrangement can be fixed with $3n-6$ coordinates. One can find these coordinates if one knows at least $4n-10$ distances between the points. Finally, there are $n(n-1)/2$ distances between the n points. The inequalities

$$n(n-1)/2 \geqslant 4n - 10 \geqslant 3n - 6 \qquad (4.41)$$

are fulfilled for $n \geq 4$. Thus, if a system consists of any number of "point-like" constituent parts, a measurement of the distances between them allows one to reconstruct their mutual spatial arrangement. In order to determine the distances, one can employ the second moments in (4.40) if one knows independently the radii of gyration of the particles that take part in the interference experiment.

Thus the mutual arrangement of 9 (out of 21) proteins of the 30S subunit of the ribosome has been reconstructed.[20]

As the number of measured distances increases, the source data for the triangulation method can be only the second moments, while the result is the coordinates and radii of gyration of the proteins in the ribosome *in situ*.[19] We should note in this regard that an additional study of the other moments of the function $p(R)$ apparently can yield certain information on the shape and relative orientation of the proteins *in situ*.

As an illustration, Fig. 1 shows two models of the 30S subunit. The upper picture is constructed from a large set of electron-microscope, biochemical, and neutron-diffraction studies without applying triangulation[21] (a); The lower diagram is constructed only on the basis of data obtained by the triangulation method[20] (b).

In closing we must take up the indirect uncertainties involved with using H–D mixtures. In addition to the inhomogeneous H–D exchange in the protein mentioned above, there is a risk of encountering substantial rearrangements of the structure of complex particles when they are put into D_2O. D_2O is well known to be toxic to higher organisms. A number of studies have noted a strong decrease in the dissociation constant of proteins into their subunits (lactate dehydrogenase[22]) and an increase in the melting point of ribosomal RNA.[23] A heightened tendency of individual proteins to aggregate formation in D_2O has been repeatedly noted. A probable mechanism of these changes is the increased strength of the hydrogen bond in a deuterated medium. In this regard, one must always make control studies in experiments employing isotope substitution to prove

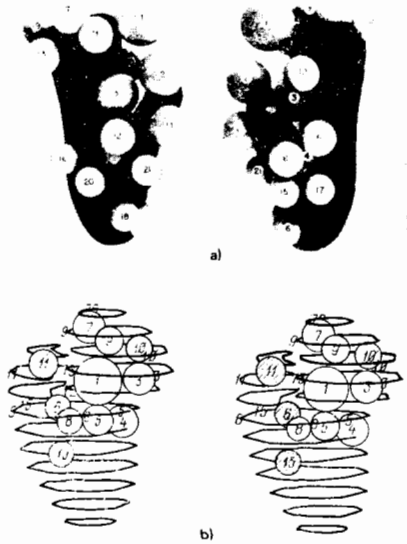

FIG. 1. Models of the 30S subunit. a) Arrangement of the 21 different proteins (balls) and the RNA shown from two sides, according to a model derived by a set of methods[21]; b) Stereo pair of the arrangement in space of 11 proteins from the results of triangulation. [Taken from: V. R. Ramakrishnan *et al.*, J. Mol. Biol. **152** (1981) with the kind permission of the authors and the publishers of Academic Press, London.]

the conservation of the structure being studied as the isotope composition of the medium is altered. One can easily carry out such a control by comparing the x-ray data for the nondeuterated and deuterated systems, since the electron density does not depend on the isotope substitution.

g) The inverse scattering problem

We denote by the term "inverse problem" the problem of reconstructing the scattering density $\rho(\mathbf{r})$ from the measured $I(\varkappa)$ relationship. A cursory glance at the fundamental formulas (3.1)–(4.1) suffices to make evident the unsolvability of this problem. In addition to the loss of the phase function $\exp[i\varphi(\varkappa)]$ [in going from (3.1) to (3.2)], a substantial fraction of the information is lost by averaging over the orientations of the macromolecules in the solution in (4.1). In this regard, fundamental application is made either to the characteristics of the macromolecule (radius of gyration, asymmetry) that have a structural meaning at a resolution of the order of the dimensions of the macromolecule, or to the distance distribution function in a homogeneous particle (4.39) or to the radial density distribution[5]:

$$W(r) = \frac{1}{2\pi^2 r} \int_0^\infty \varkappa \sqrt{I(\varkappa)} \sin(\varkappa r)\, d\varkappa. \qquad (4.42)$$

The latter has a structural meaning only for particles of spherical symmetry.

This state of affairs has led to the appearance of methods of constructing one or several models of $\rho(\mathbf{r})$ from the multitude of possible solutions. In its simplest formulation, this problem has been solved by Kratky[7] by using equivalent scattering objects. In this

method one constructs a model from one or several homogeneous objects of simple shape (ellipsoids, cylinders) such that its calculated scattering curve satisfactorily agrees with the experimental curve. The scattering curves of practically all biological macromolecules can be described by this method. Of course, this doesn't mean that all macromolecules are ellipsoids.

The Kratky method has been developed by Feĭgin and his associates[8] and it enables one to calculate the scattering curves for objects of arbitrary shape. This step has made it possible to include in the model information obtained both by small-angle scattering and by electron microscopy and other methods. Although they are not the sole solution of the inverse problem, the models of macromolecules obtained in this way at least agree with the whole set of existing experimental information.

The inverse problem has been examined in general form by Harrison and by Stuhrmann.[24] Utilizing the mathematical apparatus of spherical harmonics, this approach also enables one to synthesize a multitude of models. One cannot choose among them by physical considerations, whereas considerations of "simplicity of description" essentially return us to the equivalent ellipsoids of Kratky.

A more rigorous method of reconstructing $\rho(r)$ is the triangulation method [see (4.6)]. However, it has been applied as yet only to the ribosomal subparticles.

An expansion of the set of objects to be studied by this method seems somewhat problematical, owing to the unusual complexity of the procedure of preparing the specimens to be studied. A general key to solving the structural problem is the use of monocrystals consisting of biological macromolecules.

h) Experimental studies

The experimental methods and technique of small-angle research is the topic for a separate article, and we shall not treat them here. The aim of this section is to illustrate the approaches presented in Secs. 4c–f.

1) *Determination of compensation points, molecular weights and volumes.* Even the early studies of a protein in solution by small-angle neutron scattering that were performed on hemoglobin[25] showed that the relationship of $\sqrt{I(0)/C}$ to the fraction of D_2O in the H_2O–D_2O mixture is a straight line. This relationship is predicted by Eq. (4.27). In subsequent experiments, other biological macromolecules have also shown straight-line relationships of $\sqrt{I(0)/C}$ to the fraction of D_2O. Figure 2 shows as an example this relationship for a fragment of ribosomal RNA, for the ribosomal protein S4, and for the specific complex between them.[26] We see from this diagram that each molecule has its own compensation point. Table III summarizes the experimental data from determining the compensation points for a number of biological particles. The data of this table show the compensation points to lie close for different proteins. The mean compensation point for proteins amounts to 42% D_2O in the H_2O–D_2O mixture. This corresponds to a scattering-amplitude density of 2.36×10^{10} cm/cm^3. The small deviations from this value are correlated with the partial specific volume of the protein. The experimentally observed values of $\bar{\rho}$ at the compensation point for proteins differ by only several percent from those calculated from the chemical composition and the partial specific volume. One assumes in this calculation that the number of hydrogens replaced by deuterium equals the product of the number of all potentially exchangeable hydrogens by the fraction of D_2O in the solvent.

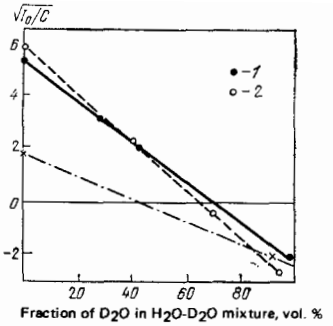

FIG. 2. Relationship of $\sqrt{I(0)/c}$ to the fraction of D_2O in the H_2O–D_2O mixture for an isolated fragment of ribosomal RNA (1), the complex (fragment of ribosomal RNA and protein S4) (2), and for the isolated protein (x).[26]

The compensation of the fragment of ribosomal RNA occurs at 70% D_2O, which corresponds to a scattering-amplitude density of 4.32×10^{10} cm/cm^3. The observed compensation points of nucleoproteins (ribosomes, viruses) lie in the interval between the compensation points of RNA and protein. To calculate the values of the compensation point of a particular ribonucleoprotein, we must know the weight fraction of one of the compenents. The influenza virus has an anomalously low experimental value of $\bar{\rho}_c$ (see Table III). This involves the presence in the virus of an appreciable fraction of phospholipids (up to 20%) having a small value of $\bar{\rho}_c$. The experimental value of the compensation point of DNA has not yet been established.

For an ensemble of heterogeneous molecules differing in scattering-amplitude density, the relationship of $\sqrt{I(0)/C}$ to the fraction of D_2O in the H_2O–D_2O mixture ceases to be linear. Experiments studying solutions of ferritin, whose molecules contained varying amounts of iron micelles inside the protein shell have indicated a considerable deviation from a straight-line relationship.[27]

The determination of the molecular weights of two-component particles at the compensation points enables one to determine the molecular weight of each of the components comprising the particle (*in situ*). This approach has been employed to study association-dissociation processes of tRNA with aminoacyl-tRNA synthetases. It has been shown for asparagine tRNA that the most stable complex is formed when one has two molecules of the synthetase per two molecules of the tRNA.[28]

2) *Determination of radii of gyration.* At present the determination of the radii of gyration R_g of macromolecules by neutron scattering has become a routine procedure. The range of R_g values accessible to this method is fixed by the specific geometry of the spectometer. Thus, e.g., the small-angle camera D11 of the spectrometer of the Laue-Langevin Institute (Grenoble) enables one to measure R_g over a range from several hundred to units of Å. Practically the same range of R_g is accessible to the method of diffuse x-ray scattering.[7]

The range of concentrations usually required for reliable measurements of R_g in light water and extrapolation to zero concentration depends on the molecular weight of the macromolecule. For proteins of molecular weights of $10-20\times10^3$ daltons, it amounts to 3–10 mg/ml. One can substantially decrease the required protein concentration by going to measurements in heavy water. Then one can reduce the protein concentration to 0.8–1 mg/ml.[29] The possibility of making such measurements is important in principle for particles of low solubility. They include, e.g., the ribosomal proteins. A measurement of their radii of gyration led the authors of Ref. 29 to conclude that most ribosomal proteins in solution amount to ordinary compact particles, neither extended nor unfolded.

The possibility of measuring the dimensions of a two-component particle at the two compensation points where each of the components successively becomes "invisible" opens up to an experimental pathway for studying the conformational changes in the components upon forming such a particle. Let us give two examples of such studies. The left-hand side of Fig. 3 shows the variation of the neutron-scattering intensity in the Guinier region for a fragment of ribosomal RNA at different contrasts. The radius of gyration proved not to depend on the contrast, and was 47.5 ± 2 Å. The right-hand side of Fig. 3 shows the analogous variation for the complex of this same fragment with the ribosomal protein S4 at various contrast. The radius of gyration proved to depend strongly on the contrast: 17–18 Å in 79% D_2O and 58 Å in 98% D_2O. A calculation by Eq. (4.31) gave the values $R_g = 50\pm3$ Å and $R_g = 17\pm3$ Å respectively for the fragment of RNA and protein S4 contained in the complex. Within the limits of experimental error these values agreed with the values of R_g of the isolated RNA fragment (47.5 ± 2 Å) and of the free ribosomal protein S4 in solution (18 ± 2 Å). Hence it was concluded that the compactness of the RNA fragment and of ribosomal protein S4 does not change upon forming the complex.[26]

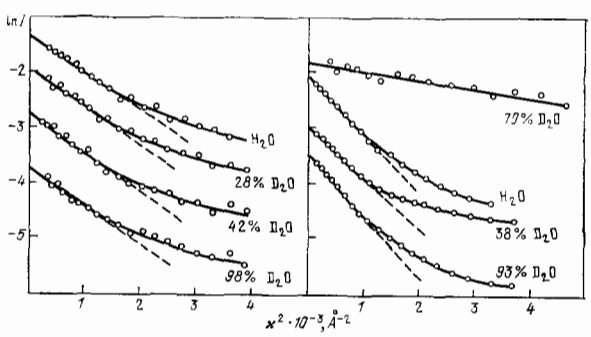

FIG. 3. Relationship of the scattering intensity to the length $\varkappa$ of the scattering vector in Guinier coordinates at different contrasts. a) Fragment of ribosomal RNA; b) complex (fragment of ribosomal RNA and protein S4).[26]

A second example involves the study of the interaction of tRNA with synthetases. It was found that the formation of the specific complex between them is accompanied by a considerable change in the radius of gyration of the protein and in the intensity extrapolated to $\varkappa = 0$.[28] The authors discuss four possible mechanisms of rearrangement in the system that would give rise to the stated changes.

3) *Study of large-scale inhomogeneities in a particle.* Many biological particles consist of two components that strongly differ in scattering-amplitude density (see Table III). If each of the components occupies a sufficiently extended region in the particle, one can establish the spatial distribution of the components by studying the dependence of R_g^2 on the contrast (fraction of D_2O in the H_2O–D_2O mixture). This approach has been widely applied for studying ribosomes, viruses, nucleosomes, and other particles.

Figure 4 shows the relationship of R_g^2 to the fraction x_p of scattering by the protein for ribosomal particles. We see here that the RNA and the protein are distributed substantially inhomogeneously in the large and the small subparticles of ribosomes. The RNA is concentrated preferentially at the center of the ribosomal particles, while the protein lies at the periphery. The distance between the centers of gravity of the RNA and the protein in the ribosomal particles is small (no more than 30 Å). This spatial arrangement of the RNA and protein in ribosomes has been observed in a number of laboratories. According to modern views, it reflects

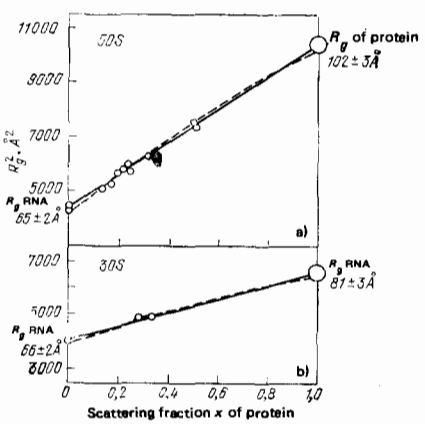

FIG. 4. Relationship of the square of the experimental radius of gyration R_g^2 to the scattering fraction x of the protein for the 50S (a) and 30S (b) ribosomal subunits.[30] The radii of gyration of RNA and protein in both subunits are calculated for $x_{protein} = 0$ and $x_{protein} = 1$, respectively. The maximum distance between the centers of gravity of RNA and protein as calculated from the curvature of the parabola (dotted line) drawn through the experimental points does not exceed 30 Å for the two ribosomal subunits.

the skeletal role of RNA in formation of the structure of the ribosomal particles.[30]

In contrast to the ribosome, it has been established by the contrast-variation method for the nucleosome (the smallest subunit of chromatin, which consists of 140 nucleotide base pairs and eight histone proteins) that the proteins lie inside the particle and the DNA outside.[31]

Application of the contrast-variation method for studying viruses has made it possible not only to confirm the well-known spatial structure of most viruses (location of RNA inside the protein shell), but also to establish new essential details of their structural organization. Thus, it has been established for several spherical viruses that the RNA and protein have an extended contact zone. Here the part of the protein globule penetrating into the RNA has a disordered conformation.[32] As the authors of Ref. 32 see it, this disorder explains the poor resolution of the structure of the inner part of a virus observed in x-ray structural studies.

Study of the globular proteins by the contrast-variation method has shown that these proteins cannot be treated in neutron scattering as objects of uniform density. A dependence of R_g^2 on the contrast has been found for lysozyme[33] and myoglobin.[15] The parameter α in Eq. (4.33) proved to be positive. This fact is interpreted as confirming the concept of a hydrophobic core in globular proteins, since the scattering-amplitude density of the hydrophobic groups is smaller than for the hydrophilic groups.

4) *Shape analysis and construction of homogeneous models.* The experimental data that we have discussed show that biological macromolecules look like inhomogeneous particles to neutrons. This situation strongly hinders the interpretation of the experimental scattering curve in the homogeneous object approximation, and it promotes to front rank the approaches that distinguish the effects of shape and of inhomogeneity. Section 4b has presented a general approach to this problem. In applying it to real molecules, difficulties in interpreting the function I_F arose, as mentioned in Sec. 4b, involving inhomogeneous exchange. The shape function I_γ has been used to construct models of lysozyme,[33] fibrinogen,[34] and the 50S subunit of the ribosome[35] by the method of spherical harmonics.[24] A model of the IgG antibody has been proposed from a set of varied data.[36] The value of these models is problematical, while the limited possibilities of the method were especially clearly manifested in the example of lysozyme, for which the shape of the molecule known from x-ray structural analysis of single crystals cannot be reconstructed by modeling.

5) *Determination of distances labeled parts of a molecule.* Binding by active centers of bivalent antibodies to a substance whose scattering-amplitude density differs from that of the antibodies enables one to realize in practice the model case of interference of two objects lying at a fixed distance from one another [see Eq. (4.35)]. This approach has been employed to determine

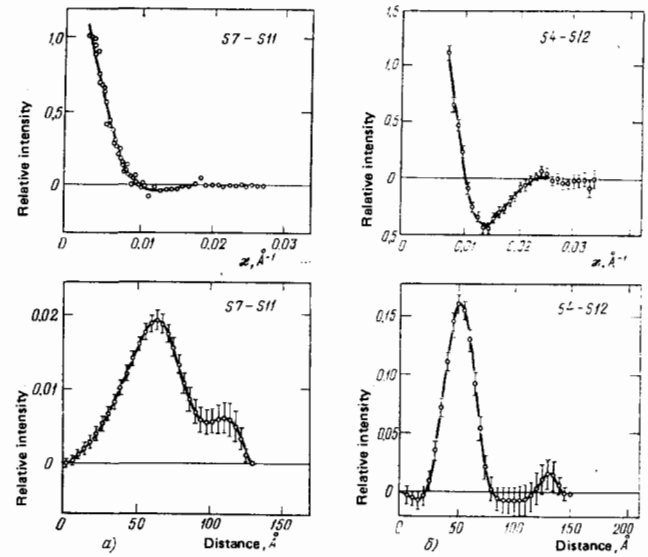

FIG. 5. Observed interference patterns (upper parts of the diagrams) for two pairs of ribosomal proteins S7–S11 (a) and S4–S12 (b), together with the distance distribution functions (lower part of the diagrams) for these proteins as calculated from these data by Eq. (4.39). The interference patterns for the pair of proteins S7–S11 are taken from: V. R. Ramakrishnan *et al.*, J. Mol. Biol. **152** (1981) with the kind permission of the authors and the publishers of Academic Press, London. The interference patterns for the pair of proteins S4–S12 are taken from: D. G. Schindler *et al.*, J. Mol. Biol. **134**, 595 (1979) with the kind permission of the authors and the publishers of Academic Press, London.

distances between the active centers of IgG immunoglobulins isolated at an early stage of the immune reaction. Two sufficiently large spherical antigens (dextran of molecular weight 40,000 with an attached dinitrophenyl group were combined with the antibody. The interference pattern of oscillations was studied in a H_2O–D_2O mixture (41%) in which the scattering from the antibody was eliminated. One of the interference maxima was found and the distance between the active centers was determined to be 276 ± 15 Å. This value was subsequently confirmed by studying the contrast-dependence of the radius of gyration.[36]

The more general problem of determining distances between the centers of gravity of labeled regions of a complex system was solved in the course of reconstructing the quaternary structure of the 30S ribosomal subunit. The experimental arrangement was explained in Sec. 4b. Figure 5 illustrates the interference pattern (upper part of the diagram) and the radial distribution functions calculated by Eq. (4.39) (lower part of the diagram) for two pairs of ribosomal proteins.

5. NEUTRON DIFFRACTION OF PROTEIN MONOCRYSTALS AND OTHER PERIODIC SYSTEMS

Structural neutron diffraction at the atomic level has traversed the path of development from the observation of the first Bragg reflections (E. Fermi and L. Marshall[37]) to the study of the structure of the most complicated systems: protein crystals. The development of

the experimental methods and of the potentialities in neutron physics has always been determined by the available power of neutron sources. At the end of the sixties these potentialities had been developed and realized to the point that successful attempts were made to apply neutrons for structural analysis of monocrystals of biological macromolecules.[38] The interest in this field was based on a number of advantages that neutron studies have over x-ray studies. They include the "sensitivity" of neutrons, which substantially differs from x-rays, to the presence in the structure of light atoms, and primarily to hydrogen; the unique sensitivity of neutrons to isotopic substitution of atoms, which is entirely lacking for x-rays; and the possibility in principle of solving the phase problem by using one isomorphous derivative of a protein containing atoms having anomalous dispersion of neutrons. Finally, in neutron irradiation the radiation load on the crystal proves to be 5–6 orders of magnitude smaller than for x-rays. This allows one to neglect in practice the problem of radiation damage to the crystal during the total time for taking the diffraction data. At the same time, the neutron method of study possesses certain defects. We should include among them the relatively high level of incoherent scattering of neutrons in hydrogen-containing materials and the presence of an appreciable inelastic scattering of neutrons. These two factors give rise to a background scattering that impedes the measurement of the intensities of the diffraction reflections. However, the fundamental "defect" of neutrons as yet is the low intensity and restricted accessibility of neutron beams, which arise from the technical capabilities of contemporary neutron sources and their high cost. Investigators "rescue" the situation both by employing monocrystals of record-making (for proteins) size (up to 30 mm^3) and by extremely long exposures (several thousand hours).

The methods of structural neutron diffraction of proteins applied up to now differ but little from the known x-ray methods, while their theoretical descriptions practically totally coincide. Consequently we shall refer interested readers to the recent monographs[39] for detailed descriptions of the theory and methodological problems. Here we shall confine the treatment to recalling several terms that we shall be employing. The *resolution d* defines the smallest distance at which one can distinguish two objects (e.g., atoms) having about the same scattering amplitudes In practice the resolution is characterized by the smallest interplanar distance d_{min} above which one has measured the intensities of Bragg reflections and included them in the treatment. The "high-resolution" region is arbitrarily defined as $d \leq 2$ Å, i.e., the region in which individual atoms are already distinguished in the Fourier maps The *phase problem* involves the fact that one needs the complex coefficients (structure factors) for a Fourier synthesis of the studied structure in real space, while experiment allows one to measure only the square of the modulus of these coefficients. The solution of the phase problem in x-ray structural analysis is based on the method of isomorphous substitution, in which one must study the protein of interest and several (3 or 4) of its derivatives with heavy atoms (mercury, gold). The *R-factor*, or discrepancy factor, is defined as

$$R = \sum ||F_0| - |F_c|| / \sum |F_0|. \qquad (5.1)$$

Here $|F_0|$ and $|F_c|$ are the measured and calculated moduli of the structure factors. It is employed as an average measure of the discrepancy between the structural model and the experimental data. Here the summation is performed over all the reflections measured and adopted for processing.

a) The problems of neutron studies

In view of the difficulties of strucutral neutron diffraction of protein crystals noted above, it is evident that neutron studies are expedient for solving problems that the x-ray method cannot handle. They include first of all the direct localization of hydrogen atoms and water. With single exceptions (e.g., Ref. 40), high-resolution x-ray studies determine the coordinates of carbon, nitrogen, oxygen, and heavier atoms. The hydrogen atoms are either included in calculating the structure factors in the final phase of structure analysis, starting with independent information on the stereochemistry of the amino-acid residues, the hydrogen bond, and the structure of water molecules, or they are not considered at all. Evidently this approach does not answer questions not solvable by using stereochemistry, and all the more so it gives no information on the true state of affairs in proteins.

The roles of hydrogen in proteins are very numerous.[41] The most crucial of them are: hydrogen of polar groups of amino-acid residues that participates in forming the structure of the protein and in enzymatic reactions; hydrogen contained in water molecules, which also determines the structure and functioning of the protein. To cast light on these questions requires neutron-diffraction studies at high resolution (better than 2 Å). Along with the laborious (and therefore few in number) experiments on the localization of hydrogen, low-resolution (10–20 Å) studies are of independent interest, in which problems of the coarser structure are solved—the mutual arrangement of homogeneous (on this scale of magnitude) components in multicomponent complexes. Such components include proteins, nucleic acids, lipids, etc. Studies of this kind are less laborious and do not require monocrystals of very high quality. For such problems neutrons prove to be an effective instrument, owing to the sufficiently great differences in scattering-amplitude densities of these components. Moreover, the contrast-variation method discussed in Sec. 4e proves to be an effective instrument also for systems of periodic structure. This enables one to identify unambiguously the contribution of the individual components to the intensity of the observed reflections, and thus to facilitate the solution of the phase problem at low resolution.

b) High-resolution studies

High-resolution neutron diffraction has been used to identify unambiguously the amino-acid residue bearing the functional group at the active center of trypsin that

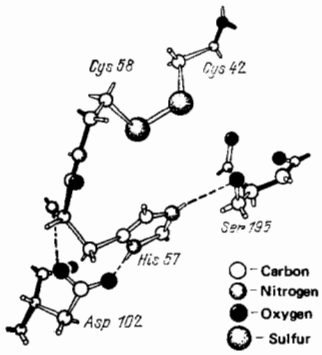

FIG. 6. Arrangement of the amino-acid residues in the active center of trypsin.[39]

acts as a chemical base during proteolysis. The earlier studies using NMR, isotope exchange, differential infrared spectroscopy, and also quantum-mechanical calculations yielded differing predictions. Up to the time of carrying out the studies to be described, the overall pattern of the proteolytic process had been reduced to the following chain of events (Fig. 6). The carbonyl carbon of the substrate is attacked by the hydroxyl oxygen of Ser-195, while simultaneously the hydroxyl proton from this group is transferred to the imidazole group of His-57. The dilemma was unsolved of whether His-57 is the final point (base) that binds the proton, or whether His-57 plays the role of an intermediary, while the base is Asp-102. To solve this problem, trypsin was inhibited with diisopropyl fluorophosphate, which binds specifically and covalently to the hydroxyl group of Ser-195 in serine proteases. The geometry of the inhibitor greatly resembles that of the expected intermediate state of the substrate in the hydrolysis reaction, where the hydroxyl proton in the enzyme has already been removed, but the peptide bond has not yet been broken. Hence a structure analysis of the trypsin-inhibitor complex should answer the question of the degree of protonation of the catalytically important amino-acid residues of the enzyme in the decisive phase of the hydrolysis. Neutron diffraction by a monocrystal of trypsin inhibited by the method stated above actually made it possible to detect unambiguously the extra proton on His-57. The experiment was performed at a resolution of 2.2 Å. The monocrystal had a volume of 1.5 mm^3 and was soaked in heavy water for a month before beginning the exposure.[42] In all, the intensities of 8700 reflections were measured. The structure was recontructed on the basis of the structural model of trypsin obtained by x-ray diffraction methods. During the reconstruction the R-factor declined from $R = 0.304$ to $R = 0.187$. In addition to the structure of the active center, the coordinates were reconstructed for about 3/4 or the total number of hydrogen atoms and 55 water molecules were localized. However, the results of the analysis of these data have not yet been published.

During the past decade the view has become generally accepted that water plays the decisive role in generating the native structure of the protein macromolecule and in the course of enzymatic processes. One of the most recent reviews on this problem[41] notes that van der Waals interactions contribute only 18–24% to the total energy of packing of the polypeptide chain into the globule. The remaining ≥75% arises from the interactions of the protein chain with water. The role of water in generating the structure of a protein is extremely complex. A classification has been proposed[41] of structured water in proteins into eight types that differ in the arrangement of the water molecules in the protein (internal and surface water) and in the type of nearest atomic environment (metal ions, charged groups, polar groups, hydrophobic groups). The list of functional roles of water is no less varied, beginning with the multitude of forms of stabilization of the protein structure (hydrophobic interactions, hydrogen bonds, water-salt bridges, shielding of internal charged ions and groups) and ending with the participation of water in enzymatic reactions (stabilization of the enzyme-substrate bond owing to displacement of water from the cavity of the active center and to the corresponding decrease in the dielectric constant, participation in processes of fast proton transport in acid-base catalysis reactions, etc.) It is difficult to overestimate the future role of neutron-diffraction studies in this field, although the existing few experiments as yet rather serve to demonstrate the complexity of the experimental approaches.

The furthest advanced are the neutron-diffraction studies of the structure of myoglobin, which have been conducted over a number of years at the Brookhaven National Laboratory.[43] The CO-derivative of sperm-whale myglobin (CO-Mb) was studied. A monocrystal of volume 24 mm^3 was soaked in heavy water for several months to replace the light water of crystallization. The data set offered a spatial resolution of 1.8 Å. The adopted procedure of replacing the light water with heavy water effectively reduced the incoherent background and thus improved the accuracy of measurement of the intensities of the reflections. However, a new problem arose here, involving the incomplete replacement of all the hydrogens with deuterium. This phenomenon strongly impedes the reconstruction of the coordinates of the atoms in real space, since a new parameter arises for each hydrogen atom: the degree of replacement. In Ref. 43 only three values were assigned to this parameter: 0, 0.5, and 1.0.

The coordinates of the oxygen and deuterium atoms of 40 water molecules were determined for the studied structure of CO-Mb. Among them, 7 molecules are bound only to other water molecules, and the rest—either to peptide atoms of the main chain or to side groups, with 27 water molecules forming bridges between atoms of the protein. The localization of water molecules in proteins is as yet ambiguous in nature. Thus, in this same protein an analysis of the neutron-diffraction data using only the x-ray coordinates for the heavy atoms revealed 106 water molecules[44]; a later x-ray study reduced this number to 72, while the most recent neutron-diffraction study,[43] which included refining the coordinates in real space, left 40 molecules, among which only 25 coincide with the x-ray positions. Perhaps a part of the disagreement involves the as yet unelucidated role of salts and of the pH in producing

protein-bound water, while another part involves individual features of the procedures employed for reconstructing the spatial structure. Probably the situation will improve as the spatial resolution is increased. A number of investigators have already reported the existence of neutron-diffraction experimental data at high resolution for CO-myoglobin ($d = 1.5$ Å),[43] ribonuclease A ($d = 2.0$ Å),[46] and hen egg-white lysozyme ($d = 1.4$ Å).[47]

c) The phase problem

As the studies discussed above have shown, the set of x-ray phases obtained without taking into account the positions of the hydrogen atoms serves as a rather good initial approximation for analyzing the neutron-diffraction data. In the course of this analysis, the refinement of the phases is an obligatory stage, and the applicability of the corresponding methods has been tested in Ref. 48. The possibility of conducting a complete structural experiment with phase determination using only one derivative of the protein is of independent interest. This derivative can contain either paramagnetic atoms or atoms having nuclei showing anomalous dispersion for neutrons (neutron resonances).[49] The scattering amplitude of neutrons in these cases becomes an experimentally variable quantity that in principle enables one to solve the phase problem. As the estimates show, if we adopt as the criterion, a 10% change in intensity upon going to the inverse reflection, it suffices to have one atom of ^{113}Cd per 1000 atoms of the protein to solve the structure. For noncentrosymmetric structures one needs measurements at two different wavelengths of neutrons. This possibility has been studied experimentally with Cd-myoglobin.[48] The corrections to the x-ray phases on the average amount to about 37°. Although there are at present no more detailed experimental studies, the theoretical study of this possibility is continuing.[49]

d) Low-resolution studies

The simplest considerations imply that the exposure time increases with increasing resolution d no more slowly than V/d^3, where V is the volume of the unit cell. Hence a relatively large number of studies have been performed with substantially poorer resolution (10–20 Å), which shortens the duration of the experiment to several days or hours. Naturally, it makes no sense here to treat the structure of the objects of study on the atomic level, but the possibility remains of distinguishing components having different mean coherent scattering-amplitude densities. In many cases such "coarsened" structural information proves to be important and interesting. Moreover, far from all biological objects can be obtained in the form of sufficiently perfect crystals. Here the method of contrast variation of the solvent that was treated in Sec. 4e for dilute solutions of macromolecules proves very effective. The study of the potentialities of this method for crystals at low resolution carried out in Ref 50 showed that, at low enough resolution ($d > 10$ Å), the structure factors of elementary volumes of the order of d^3 vary linearly with the contrast. In more developed form, this approach has been applied to study a constituent component of chromatin—the nuclear particles (nucleosomes),[51] which contain approximately equal amounts of protein and DNA. Monocrystals of volume 0.15–0.015 mm^3 were employed, having unit-cell dimensions $a = 119$ Å, $b = 198$ Å, $c = 111$ Å, symmetry $P2_12_12_1$. Data were taken out to $d = 22$ Å, which allowed measuring 114 reflections. Four different compositions were used of the mother liquor (H_2O, 39% D_2O, 65% D_2O, and 90% D_2O), in which the crystals were soaked for no less than two weeks before the exposures. The results of the data collection were subjected to the traditional procedure of crystallographic analysis: Patterson synthesis, phase determination, and reconstruction of the three-dimensional structure with subsequent refinement of the phases. The result of the analysis was a three-dimensional model of the structure of the nuclear particles in which the helical DNA is wrapped around a protein core (histone octamer). The radius of the superhelix is 49 Å, pitch 27.5 Å, and 1.8 turns of DNA were wrapped on one octamer. These results have substantially refined the concepts derived from the small-angle studies.[31]

The low-resolution approach described here has been applied to a large set of biological objects: calcification of tendons,[52] the internal organization of membranes,[50] the structure of viruses,[17] etc. It is interesting to note in this regard that in recent years magnetic fields of intensities up to 18 teslas have successfully begun to be utilized,[50,54] to prepare moderately ordered system. Thus success was achieved in orienting several types of membranes, and rod-shaped viruses, and also in obtaining oriented fibrin gels by polymerization in a magnetic field.

6. STUDIES OF THE DYNAMICS OF MACROMOLECULES

In addition to structural studies proper, the neutron-diffraction methods in principle enable one to study dynamic processes in biological systems, e.g., diffusion of small molecules (water), segmental mobility, longitudinal and transverse vibrations of polypeptide chains, etc. One of the successful steps in this direction is the study by quasielastic neutron scattering of the migration of water of hydration over the surface of the protein C-phycocyanin[55]. Deuterated protein isolated from blue-green algae and microorganisms grown in a deuterated medium was used as specimens. This measure substantially reduced the background neutron scattering from the hydrogen atoms contained in the protein. From 300 to 500 mg of the protein was kept in a helium atmosphere having a controlled content of H_2O vapor, which allowed the degree of hydration of the protein to be fixed. The measurements were performed with applied pulses in the range 0.11–0.33 Å^{-1} with an energy resolution of the order of 2×10^{-6} eV in the region of applied energies $\leq 62 \times 10^{-3}$ eV. The dependences of the broadening of the quasielastic line on the magnitude of the applied pulse that were obtained revealed an oscillating character, which indicates a jumpwise mechanism of diffusion of the water molecules. The authors estimate the mean jump length of the molecule to be 7–9 Å, and the settled lifetime (between two consecutive jumps) to be $(1.5–3.0) \times 10^{-9}$ s.

7. RESULTS AND PROSPECTS

In only 10 years the scattering of thermal neutrons has become recognized as an instrument for structural studies in molecular biology. The high-resolution methods have already made it possible to fill in the details of the views on the course of proteolytic reactions and have posed a number of new problems involving the localization of water molecules and H–D exchange. The low-resolution methods, which have become very widely known owing to their relative simplicity of experiment, have been successfully applied both for structural problems proper and for studying the changes of structure as macromolecules perform their biological functions. Apparently these methods will be widely applied also in the future for solving problems in which the "molecular" structural level suffices.

There is an intermediate region of structural organization between an ideal solution and an ideal monocrystal. Generally we should be able to obtain structural information from it with a resolution intermediate between the dimensions of the macromolecule and of an atom. The great uncertainty in the forms of structural organization in this region does not allow a general solution of the structural problem. However, whenever this uncertainty is removed, the problem becomes structure-sensitive. In particular, in principle this approach can be used not only to study the distances and arrangement of particles in specifically bound dimers and more complex aggregates (see Sec. 4f), but also to extract information on their shape, internal structure, and mutual orientation. This approach has not yet been studied in detail, neither on the theoretical level nor experimentally, though the development of the neutron-diffraction methods of distinguishing the interference functions is a beginning that promises much.

Up to now the extent of use and the accessibility of the neutron-diffraction methods have left much to be desired. In Western Europe the leading center for these studies has become the Laue-Langevin International Institute at Grenoble, France. It is equipped with a cold moderator, two instruments for studying small-angle scattering, and also a number of diffractometers and inelastic-neutron-scattering spectrometers. In the USA the Brookhaven National Laboratory has a neutron diffractometer and a small-angle-scattering apparatus. Since the reactor there is not equipped with a cold moderator, diffraction studies on monocrystals have been preferentially performed there in recent years. The National Laboratory at Oak Ridge has completed the creation of an all-nation center for small-angle studies,[56] which possesses at the same time apparatus for studying small-angle scattering of neutrons and x-rays. In the Soviet Union neutron-diffraction studies of biological macromolecules have been begun at the Laboratory of Neutron Physics Of the Joint Institute for Nuclear Research (Dubna) and the B. P. Konstantinov Leningrad Institute of Nuclear Physics (Gatchina).[36,57]

In the very near future we should expect further growth of neutron-diffraction studies. The number of operating instruments at existing reactors is increasing on an extensive level, and also new neutron sources are going into operation: the "Orpheus" reactor in France, the pulsed neutron sources IPNS in England, ZING in the USA, the PIK reactor at the Leningrad Institute of Nuclear Studies (Gatchina) and the pulsed reactor IBR-2 at the Joint Institute for Nuclear Research (Dubna). They are all being equipped with spectrometers for neutron-diffraction studies of biological macromolecules. Simultaneously with this, the developmental work directed toward improving the efficiency of utilization of neutron beams shows great prospects. For continuous neutron sources, such pathways are the improvement of monochromators and the development of two-dimensional coordinate-sensitive detectors,[58] which enable one simultaneously to detect scattered neutrons over a broad range of angles without losing angular resolution (two-dimensional data collection). Pulsed neutron sources add the possibility of employing the entire spectrum of thermal neutrons simultaneously over a broad range of angles, and thus of collecting data in a three-dimensional volume of scattering-vector space.[59] All these developments are based on widespread application of the contemporary advances in electronics and computational technology, and they should increase the efficiency of structural studies by a factor of 10^2–10^3.

Studies of the dynamics of biological macromolecules by neutron-diffraction methods is a field that has hardly been touched as yet. In part this involves the restructed potentialities of neutron technology (energy resolution of 10^{-6}–10^{-9} eV, which corresponds to characteristic times of 10^{-10}–10^{-7} s), and in part the insufficiently detailed understanding of the behavior of macromolecules in this time range. The first experiments in this field have been just begun.[60] We suppose that one should expect rapid progress in this field as well.

In closing we deem it necessary to express our sincere gratitude to A. S. Spirin and I. M. Frank for many years of support of our interest and efforts in the field of neutron-diffraction studies of biological objects.

[1]I. M. Frank, Priroda No. 9, 24 (1972).
[2]A. I. Akhiezer and I. Ya. Pomeranchuk, Nekotorye voprosy teorii yadra (Some Problems of the Theory of the Nucleus), Gostekhizdat, M., 1950; N. A. Vlasov, Neĭtrony (Neutrons), Nauka, M., 1971; I. I. Gurevich and L. V. Tarasov, Fizika Neĭtronov nizkikh énergiĭ (Low Energy Neutron Physics), Nauka, M., 1965 (Engl. Transl., North-Holland, Amsterdam, 1968); C. E. Bacon, Neutron Diffraction, 3rd edn., Clarendon Press, Oxford, 1975.
[3]C. E. Bacon, Acta Crystallogr. Sect. A **28**, 537 (1972).
[4]J. M. Cowley, Diffraction Physics, 2nd edn., North-Holland, Amsterdam, 1981 (Russ. Transl. of 1st edn., Mir, M., 1979).
[5]A. Guinier and G. Fournet, Small Angle Scattering of X-Rays, Willey and Chapman, New York, London, 1955.
[6]V. N. Tsvetkov, V. E. Éskin, and S. Ya. Frenkel', Struktura makromolekul v rastvorakh (Structure of Macromolecules in Solutions), Nauka, M., 1964.
[7]O. Kratky and I. Pilz, Quart. Rev. Biophys. **5**, 481 (1972).
[8]Yu. A. Rol'bin, R. L. Kayushina, L. A. Feĭgin, and B. M. Shedrin, Kristallografiya **18**, 701 (1973) [Sov. Phys. Crystallogr. **18**, 442 (1973)].
[9]H. Stuhrmann and R. Kirste, Z. Phys. Chem. (Frankfurt am Main) **46**, 247 (1965); H. Stuhrmann, Acta Crystallogr. Sect. A **26**, 297 (1965).
[10]B. Jacrot and G. Zaccai, Biopolymers **20**, 2413 (1981).
[11]A. A. Zamyatnin, Prog. Biophys. Mol. Biol. **24**, 107 (1972).
[12]R. Giege, G. Zaccai, B. Jacrot, and Li Long Ci, ILL Annual Report. Annex. ILL, Grenoble, 1980, p. 337.

[13]I. N. Serdyuk and B. A.Fedorov, J. Polym. Sci., Polymer Lett. **11**, 645 (1973); I. N. Serdyuk, Dokl. Akad. Nauk SSSR **217**, 232 (1974); I. N. Serdyuk and A. K. Grenader, FEBS Lett. **59**, 133 (1975).
[14]H. B. Stuhrmann and R. G. Kirste, Z. Phys. Chem. (Frankfurt am Main) **56**, 334 (1967).
[15]K. Ibel and H. B. Stuhrmann, J. Mol. Biol. **93**, 255 (1975).
[16]R. R. Crichton, D. M. Engelman, J. Haas, M. H. J. Koch, P. B. Moore, and R. Parfait, Proc. Nat. Acad. Sci. USA **74**, 5547 (1977); H. B. Osborne, C. Sardet, M. M. Willaz, and M. Chabre, J. Mol. Biol. **123**, 177 (1978).
[17]O. Kratky and W. Worthmann, Monatsheft Chem. **76**, 263 (1947); B. K. Vaĭnshteĭn, N. I. Sosfenov, and L. A. Feĭgin, Dokl. Akad. Nauk SSSR **190**, 574 (1970) [Sov. Phys. Dokl. **15**, 12 (1970)]; W. Hoppe, Isr. J. Chem. **10**, 321 (1972); D. M. Engelman and P. B. Moore, Proc. Nat. Acad. Sci. USA **69**, 1997 (1972).
[18]P. B. Moore, J. A. Langer, B. P. Schoenborn, and D. M. Engelman, J. Mol. Biol. **112**, 199 (1977).
[19]P. B. Moore and E. Weinstein, J. Appl. Crystallogr. **12**, 321 (1979).
[20]D. G. Schindler, J. A. Langer, D. M. Engelman, and P. B. Moore, J. Mol. Biol. **134**, 595 (1979).
[21]A. S. Spirin, I. N. Serdyuk, I. L. Shpungin, and V. D. Vasil'ev, Mol. Biol. **6**, 1384 (1979).
[22]R. F. Henderson and T. R. Henderson, J. Biol. Chem. **245**, 373 (1970).
[23]S. Lewin and B. A. Williams, Arch. Biochem. Biophys. **144**, 1 (1971).
[24]S. C. Harrison, J. Mol. Biol. 42,457 (1969); H. B. Stuhrmann, Acta Crystallogr. Sect. A **26**, 297 (1970).
[25]J. Schelten, P. Schlecht, W. Schmatz, and A. Mayer, J. Biol. Chem. **247**, 5436 (1972).
[26]I. N. Serdyuk, J. L. Schpungin, and G. Zaccai, J. Mol. Biol. **137**, 109 (1980).
[27]H. B. Stuhrmann, J. Appl. Crystallogr. **7**, 173 (1974).
[28]P. Dessen, S. Blanquet, G. Zaccai, and B. Jacrot, J. Mol. Biol. **126**, 293 (1978); G. Zaccai, P. Morin, B. Jacrot, D. Moras, J. C. Thierry, and R. Giege, J. Mol. Biol. **129**, 483 (1979).
[29]I. N. Serdyuk, G. Zaccai, and A. S. Spirin, FEBS Lett. **94**, 349 (1978).
[30]I. N. Serdyuk, A. K. Grenader, and G. Zaccai, J. Mol. Biol. **135**, 691 (1979).
[31]E. M. Bradbury, J. P. Baldwin, B. G. Carpenter, R. P. Hjelm, R. Nancock, and K. Ibel, Brookhaven Symp. Biol. **27**, IV 97 (1975); J. F. Pardon, D. L. Worcester, J. C. Wooley, K. Tatchell, K. E. Van Holde, and B. M. Richards, Nucleic Acids Res. **2**, 2163 (1975).
[32]C. Chauvin, J. Witz, and B. Jacrot, J. Mol. Biol. **124**, 614 (1978).
[33]H. B. Stuhrmann and H. Fuess, Acta Crystallogr. Sect. A **32**, 67 (1976).
[34]G. Marguerie and H. B. Stuhrmann, J. Mol. Biol. **102**, 143 (1976).
[35]H. B. Stuhrmann, M. H. J. Koch, R. Parfait, J. Haas, K. Ibel, and R. R. Crichton, Proc. Nat. Acad. Sci. USA **74**, 2316 (1977).
[36]L. Cser, F. Franěk, I. A. Gladkikh, R. S. Nezlin, J. Novotný, and Yu. M. Ostanevich, FEBS Lett. **93**, 312 (1978); Immunol. Lett. **1**, 185 (1980); Eur. J. Biochem. **116**, 109 (1981).
[37]E. Fermi and L. Marshall, Phys. Rev. **71**, 666 (1947).
[38]F. M. Moore, B. T. M. Willis, and D. C. Hodgkin, Nature **214**, 130 (1967); B. P. Schoenborn, Nature **224**, 143 (1969).
[39]T. L. Blundell and L. N. Johnson, Kristallografiya belka (Protein Crystallography), Mir, M., 1979.
[40]B. W. Dijkstra, K. H. Kalk, W. G. J. Hol, and J. Drenth, J. Mol. Biol. **147**, 97 (1981).
[41]J. L. Finney, in: Water. A Comprehensive Treatise, ed. F. Franks, Plenum Press, New York, 1979, Vol. 6, p. 47.
[42]A. A. Kosiakoff and S. A. Spencer, Nature **288**, 414 (1980).
[43]B. P. Schoenborn and J. C. Hanson, in: Water in Polymers, ed. S. P. Powland, Am. Chem. Soc. Symp, 1980, Vol. 127, p. 215; J. C. Hanson and B. P. Schoenborn, J. Mol. Biol. **153**, 117 (1981).
[44]B. P. Schoenborn, Cold Spring Harbor Symp. Quant. Biol. **36**, 569 (1971).
[45]A. Takano, J. Mol. Biol. **110**, 537 (1977).
[46]A. Wlodawer, Acta Crystallogr. Sect. B **36**, 1826 (1980).
[47]G. A. Bentley and S. A. Mason, Phil. Trans. Roy. Soc. London Ser. B **290**, 505 (1980).
[48]B. P. Schoenborn and R. Diamond, in: Proc. Brookhaven Symposium in Biology, Brookhavenm 1975/P II-3.
[49]S. Ramaseshan, Curr. Sci. India **35**, 87 (1966); A. K. Singh and S. Ramaseshan, Acta Crystallogr. Sect. B **24**, 35 (1968); F. L. Shapiro, Probl. Fiz. ÉChAYa **2**, 975 (1972); W. Jauch and H. Dachs, Acta Crystallogr. Sect. A **31**, 162 (1981).
[50]H. Saibil, M. Chabre, and D. Worcester, Nature **262**, 266 (1976).
[51]G. A. Bentley, J. T. Finch, and A. Lewit-Bentley, J. Mol. Biol. **145**, 771 (1981).
[52]S. W. White, D. J. S. Hulmes, A. Miller, and P. A. Timmins, Nature **266**, 421 (1977).
[53]B. Jacrot, Comprehensive Virology **17**, 129 (1980).
[54]J. Torbet, J. M. Freyssinet, and G. Hudry-Clergeon, Nature **289**, 91 (1981).
[55]H. D. Middendorf and J. Randall, Phil. Trans. Roy. Soc. London Ser. B. **290**, 639 (1980).
[56]W. C. Koehler, Oak Ridge Nat. Lab. Review, Summer 1980, p. 31.
[57]M. M. Agamalyan, G. M. Drabkin, A. A. Dovzhikov, *et al.*, Kristallografiya **27**, 92 (1982).
[58]T. Springer, Phil. Trans. Roy. Soc. London Ser. B **290**, 673 (1980); D. M. Kheĭker, Kristallografiya **23**, 1288 (1978) [Sov. Phys. Crystallogr. **23**, 729 (1978)].
[59]A. M. Balagurov, V. E. Novozhilov, Yu. M. Ostanevich, and V. D. Shibaev, Soobshch. OIYaI R14-12840, Dubna, 1979; A. M. Balagurov, V. I. Gordelii, M. Z. Ishmukhametov, V. E. Novozhilov, Yu. M. Ostanevich, B. N. Savenko, and V. D. Shibaev, Soobshch. OIYaI R14-80-440, Dubna, 1980; S. W. Peterson, A. H. Reis, A. J. Schultz, and P. Day, Adv. Chem. Series No. 186, p. 75 (1980).
[60]J. Randall and H. D. Middendorf, see Ref. 12, p. 346; Y. Alpert, *ibid*, P. 347.
[61]H. B. Stuhrmann, J. Haas, K. Ibel, M. H. J. Koch, and R. R. Crichton, J. Mol. Biol. **100**, 399 (1976).
[62]P. B. Moore, D. M. Engelman, and B. P. Schoenborn, Proc. Nat. Acad. Sci. USA **71**, 172 (1974).
[63]P. Beadry, H. V. Peterson, M. Grunberg-Manago, and B. Jacrot, Biochem. Biophys. Res. Commun. **72**, 391 (1976).
[64]H. B. Stuhrmann, J. Haas, K. Ibel, B. De Wolf, M. H. J. Koch, R. Parfait, and R. R. Crichton, Proc. Nat. Acad. Sci. USA **73**, 2379 (1976).
[65]H. B. Stuhrman, M. H. J. Koch, R. Parfait, J. Haas, K. Ibel, and R. R. Crichton, J. Mol. Biol. **119**, 203 (1978).
[66]S. J. Perkins, A. Miller, H. Weiss, and K. Leonard, see Ref. 12, p. 351.
[67]J. E. Mellema, S. Cusack, and A. Miller, *ibid.*, p. 358.
[68]S. Cusack, A. Miller, P. C. Krijgsmann, and J. E. Mellema, *ibid.*, p. 360.
[69]P. Boulanger, C. Devaux, and B. Jacrot, *ibid.*, p. 361.
[70]S. J. Perkins, A. Miller, *et al.*, *ibid.*, p. 340.
[71]G. M. Freyssinet and G. Tourbet, *ibid.*, p. 350.
[72]H. B. Stuhrmann, A. Tardieu, L. Mateu, C. Sardet, V. Luzzati, L. Aggerbeck, and A. M. Scanu, Proc. Nat. Acad. Sci. USA **72**, 2270 (1975).
[73]R. P. Hjelm, Jr., J. P. Baldwin, and E. M. Bradbury, in: Methods in Cell Biology, eds. G. Stein, J. Stein, and P. Kleinsmith, Academic Press, New York, 1978, Vol. 18, p. 295.
[74]B. Jacrot, Rep. Prog. Phys. **39**, 911 (1976).

Translated by M. V. King